Title
Prediction of Abdominal Aortic Aneurysm Growth by Automatic Segmentation and Radiomics Feature Quantification

Author
Xiong, Fei

Publication Date
2019

Peer reviewed|Thesis/dissertation

Prediction of Abdominal Aortic Aneurysm Growth by Automatic Segmentation and Radiomics Feature Quantification

by
Fei Xiong

THESIS
Submitted in partial satisfaction of the requirements for degree of
MASTER OF SCIENCE

in

Biomedical Imaging

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
David Saloner
F1B602CAB7874F1...
Chair

DocuSigned by:
David Saloner
DocuSigned by:
Susan Noworolski
DocuSigned by:
Jing Liu
DocuSigned by:
Joseph Leach
D80E9C46086A4BF...
Committee Members

Acknowledgements

I would like to express sincere gratitude to my advisor Dr. David Saloner for the continuous support of my M.S. study and related research, for his patience, guidance, and knowledge. I also thank Dr. Joseph Leach for providing a large set of CT images. I am also grateful to the other two members of my thesis committee, Dr. Jing Liu and Dr. Susan Noworolski who provided insight and expertise that greatly assisted the research. I would like to give my special thanks to our lab members, Dr. Yan Wang, Dr. Dimitrios Mitsouras, and Dr. Evan Kao, for inspiring me to explore new knowledge and to dive into radiomics and machine learning. Thanks to my fiancée, for your trust and love.

Abstract

Prediction of Abdominal Aortic Aneurysm Growth

by Automatic Segmentation and Radiomics Feature Quantification

Fei Xiong

An accurate assessment of abdominal aortic aneurysm (AAA) progression is essential to its clinical management. Currently, the maximum diameter of AAA at diagnosis is considered as the primary indicator of rupture risk. However, it is not optimal as rupture can happen at any size. Several patient-specific factors may also influence AAA rupture risk. Given the clinical variability in aneurysm progression, additional prognostic markers are desirable to enhance patient-specific risk stratification. Radiomics is an image processing technique that extracts quantitative and high-dimensional features from medical images. While it has emerged as a novel approach for solving diagnosis in oncology, its application in cardiovascular diseases is still limited.

This study set out with an aim to determine the feasibility of radiomics in identifying AAA with a fast growth rate (≥0.3cm/year) using CT images. An automatic AAA segmentation algorithm was developed in our pipeline. Based on the radiomics features of an 84 CT dataset, supervised classification models were implemented with two feature selection algorithms and two classifiers in a machine-learning framework. An AUC of 0.80 was achieved and the predictive power was proved through comparisons to the maximum diameter and conventional risk factors. Further multivariate analysis suggested that a radiomics-based classification model could be used as an independent, yet strong predictor for fast AAA growth rate.

Table of Contents

List of Figures

List of Tables

1. INTRODUCTION

1.1 Clinical Background

An abdominal aortic aneurysm (AAA) is a focal, balloon-like dilation of the abdominal aorta with maximum diameter exceeding 50% of its original size or greater than 3cm [1]. It affects up to 8% of men over age 65 and is becoming increasingly common in women [2, 3]. Often being asymptomatic, fast-growing AAA without intervention can lead to rupture. Given a mortality rate of 65-85%, ruptured AAA is now ranked as the 15th leading cause of death in the United States [4]. Monitoring the disease progression is therefore essential to its clinical management.

Currently, the maximal diameter (D_{max}) remains the primary indicator for AAA growth. Surveillance continues until the aortic diameter approaches 5.5cm [4], at which point surgical treatment by open or endovascular aneurysm repair is undertaken. Meta-analysis has revealed that AAA can rupture at any size before achieving the intervention criteria and thus require further considerations [5, 6]. Fast AAA growth rate (>1cm/year) is considered another important indicator for surgery[1, 7]. Nonetheless, extensive systematic reviews found that the average growth rate for AAA over 4.5cm in diameter is around 0.5cm/year [8], which is much lower than the threshold for intervention. Moreover, a number of patient-specific factors may also influence AAA rupture risk. Smoking was found to be associated with faster aneurysm growth rates, whereas increased BMI, diabetes, and hypertension were associated with slower growth rates [8]. Although the aneurysm size is regarded as the best predictor for rapid growth rate, clinical decision-making can be complicated by the other factors. Therefore, identification of fast-growing AAA or impending rupture remains challenging [9].

1.2 Imaging Markers

A wide range of imaging markers was explored to predict growth of AAA based on computed tomographic angiography (CTA). With three-dimensional (3D) visualization of the aorta, CTA provides additional information regarding the presence of intraluminal thrombus (ILT), calcified deposits, and involvement of visceral arteries [1, 10]. Independent of initial AAA diameter, the volume of ILT was found to be highly associated with rapid expansion [11, 12]. While ILT is regarded as a protective factor against large wall tension[4, 13, 14], it also serves as the inflammatory nidus of proteolysis activities that weaken the adjacent arterial wall [15, 16]. A high degree of calcification has long been treated as a potential risk factor for AAA rupture [17], which was found to adversely increase the maximum wall tension and which alters stress distribution in a finite element analysis [18]. In a retrospective study of 76 AAAs, a classification model built based on maximum diameter, surface area, tortuosity and ILT volume ratio produced a prediction accuracy of 86.6% [19]. Another classifier built with geometric indices derived from local diameter and wall thickness also has a reported accuracy of 95.6% to identify cases that need surgical repair [20]. Despite the potential impacts on AAA rupture, the prognostic value of these markers is still controversial and not determined collectively. This motivates us to explore the potential use of radiomics features as a surrogate marker in identification of fast-growing AAA.

1.3 About Radiomics

Radiomics is an image processing technique that extracts quantitative and high-dimensional features from medical images [21]. Beyond the ability to measure 3D-shape metrics, radiomics offer comprehensive characterizations of tissue via intensity-based

metrics, like texture and heterogeneity [22]. Along with the use of machine learning techniques, it allows researchers to retrieve valuable clinical information from images that is invisible to the naked eye [23]. Texture analysis on CT images was successfully used for the prediction of endovascular leak after endovascular prostheses [24]. The CT signal heterogeneity in AAA has also been shown to correlate with the expansion of small aneurysms and has been suggested to be a risk stratification tool [25]. Moreover, radiomics features have shown significant sensitivity in detection of napkin ring sign in coronary plaque so as to predict a major adverse cardiac event[26]. From these promising results, we hypothesize that radiomics features in AAA have predictive values for fast growth rate and can serve as surrogate markers for rupture.

1.4 Main Objective

The present study aims to determine the feasibility of radiomics in identifying AAA with a fast growth rate (≥0.3cm/year) using CT images. To achieve this goal, a radiomics pipeline was developed with three major components. First, an automatic segmentation algorithm was proposed to extract the volume of interest in AAA. Second, two feature selection algorithms and classifiers were implemented to determine an optimal machine-learning (ML) framework for AAA classification. Lastly, the performance of the radiomics classification model was evaluated via comparisons to the conventional risk factors, including D_{max}, smoking, age, BMI, diabetes mellitus (DM), and hypertension (HT).

2. METHODOLOGY

2.1 Study Population

The CT dataset was collected retrospectively from 230 patients at San Francisco Veterans Affairs Medical Center. All image acquisitions were performed with helical CT/MDCT scanners (mostly GE) using standard clinical protocols. Routine portal-venous-phase CT and multiphase CT evaluations were both included in the study. As part of the enrollment criteria, the patient needed to have at least one follow-up scan six months after an initial diagnosis of the disease (2 scans in total). Information pertaining to age, gender, smoking history, medications, and diagnostic outcome was available for all included patients. The exclusion criteria were post-open surgery, post-endovascular aortic repair, and poor contrast enhancement of AAA. In addition, CT datasets with slice thickness larger than 5mm were excluded from the study to ensure reasonable image quality for radiomics analysis. Finally, 84 patients were enrolled in this study. All selected CT images (512x512) have spatial resolutions within 0.459-0.976mm and were reconstructed at 1.25- to 5mm-thickness. The demographic information of patients is summarized in Table 1.

2.2 Image Analysis

The CT images were transferred in Digital Imaging and Communications in Medicine (DICOM) format to a medical image viewer (Horos, version 3.0) and reviewed by two trained radiologists with 6- and 8- year experience. Using a multi-planner reconstruction (MPR) method, the D_{max} was measured for each AAA at the baseline and follow-up time-points. Subsequently, the annual growth rate was calculated from the two measurements. Using the upper quartile (0.3 cm/year) of the current database as a divider, the patients were stratified into fast-growing (n=44) and stable (n=40) groups.

The inter-reader agreement for all measurements were assessed and verified in a previous study by our team. In addition to D_{max}, other conventional risk factors including age, BMI, smoking, DM, HT were tabulated for prediction performance analysis.

Table 1. The characteristics of patient-specific risk factors associated with the fast-growing and stable groups.

Characteristics (Mean ± SD)	**All** (n=84)	**Fast-Growing** (n=44)	**Stable** (n=40)
Gender	M	M	M
Age (years)	72.5 ± 9.4	72.6 ± 9.3	72.4 ± 9.5
Smoking (n)	27	16	11
BMI (Kg/m^2)	27.1 ± 4.6	26.0 ± 4.0	28.3 ± 5.0
Diabetes Mellitus (DM) (n)	14	5	9
Hypertension (HT) (n)	69	35	34
Intraluminal Thrombus (ILT) (n)	51	33	18
Maximal Diameter (D_{max}) (cm)	4.3 ± 1.0	4.6 ± 1.0	3.9 ± 0.76
Growth Rate (cm/year)	0.3 ± 0.2	0.4 ± 0.2	0.1 ± 0.1

2.3 Radiomics Overview

The radiomics analysis (Figure 1) began with automatic AAA segmentations in MATLAB (version R2018b). From the VOI-masked CT images, the radiomics features were extracted using Pyradiomics [27] in Python (version 3.6). A machine-learning (ML) framework was subsequently built with two feature selection algorithms and two classifiers using scikit-learn[28]. The prediction accuracies and receiver operating characteristic (ROC) curves were first evaluated in the ML framework with cross-

validation. Lastly, the statistical analysis was conducted in STATA to further examine the significance of the radiomics classification model.

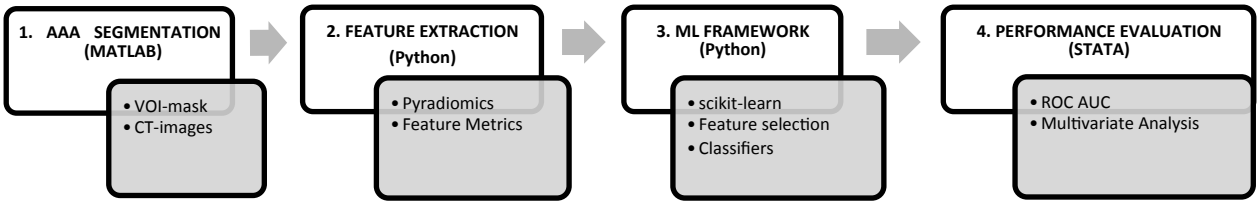

Figure 1. The radiomics pipeline for AAA classification in CT images.

2.4 Semi-automatic AAA Segmentation

An accurate and reproducible segmentation algorithm is essential for subsequent radiomics analysis. Existing methods for AAA segmentation rely on either a manual process or 3D modeling software, which are time consuming and impractical for radiomics. As the outer wall of AAA is a diffuse object with low contrast relative to neighboring tissue, a precise delineation remains difficult in CT images. To overcome such challenges, a semi-automatic segmentation algorithm was implemented with a novel geodesic active contour (GAC) model. Using an implicit representation of the contour and an added elliptical shape constraint, the segmentation algorithm effectively preserves the aortic wall integrity while isolating it from the confounding structures. It further enables a user-interactive environment in MATALB, which allows for necessary interventions to achieve better precision. A detailed mathematical explanation of the GAC model is provided in the following section.

2.4.1 Segmentation workflow

Figure 2 illustrates the workflow for the semi-automatic segmentation process. Before AAA segmentation, the axial CT images were cropped to a field of view (FOV)

that focalizes the abdominal aorta. To ensure consistency, all the segmentations started from a proximal slice near to the renal branches and ended at a distal slice at the iliac bifurcation site. A thresholding step was done to select the lumen region and replace it with the mean intensity of the surrounding tissues. This procedure helped to eliminate undesired contrast that may affect the segmentation performance. To initiate the GAC evolution, a reference contour was defined by manually delineating the outer wall of the AAA in a selected middle slice. Through an iterative process, the contour evolved adaptively to an optimal shape according to a computed deformation force. It continued until a steady-state condition was satisfied or maximum iteration was reached. Utilizing the result from the previous slice, the segmentation algorithm automatically detected the contours in other slices efficiently and generated a VOI for the AAA.

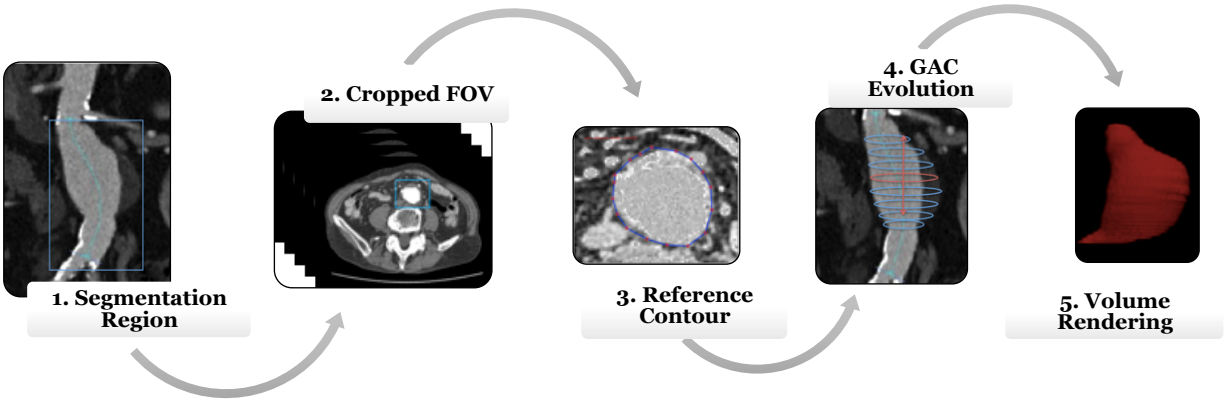

Figure 2. The semi-automatic segmentation workflow for AAA outer wall.

2.4.2 Geodesic Active Contour (GAC)

2.4.2.1 General Definition

GAC is defined as an energy minimizing spline that is associated with internal and external constraint forces [29], in which the former is a spline force that controls the smoothness of the evolving curve, whereas the latter is an image force that pushes the

curve towards the edge object. Embedded with a level-set function (LSF)[29, 30], the active contour can be prescribed as the zero level-set of a higher dimensional function. This implicit representation allows numerical computation to be performed on a fixed Cartesian grid without the need to parameterize the points on a curve. Complex topological changes like splitting or merging events can be tracked accurately.

2.4.2.2 Mathematical Algorithm

Let I be a given image domain $\Omega \rightarrow \boldsymbol{R}$, the GAC is denoted as $C(x,t) - \boldsymbol{R^2}$ and represented as the zero level-set of LSF $\Phi(x,t) \rightarrow \boldsymbol{R}$. A signed distance function is normally used for LSF, which is negative inside the zero level-set domain and positive outside. The mathematical equation is given as,

$$\Phi(x,t) = \begin{cases} -d & x \in \Omega^- \\ +d & x \in \Omega^+ \\ 0 & x \in C \end{cases} \qquad \text{Equation 1,}$$

where d is the Euclidian distance of evolving curve C to the initial curve C_0. The speed of contour motion is specified as a scalar function F of the curvature $K = div\,(\frac{\nabla\Phi}{|\nabla\Phi|})$:

$$\begin{cases} \Phi(x,t) = -F(K)|\nabla\Phi| \\ \Phi(C_0(x,t),0) = 0 \end{cases} \qquad \text{Equation 2,}$$

where ∇ is a gradient operator. For simplicity, $\Phi(x,t)$ is abbreviated as Φ in subsequent equations. As classic GAC is known to develop irregularities during its evolution, a new variational level set formulation is implemented in our segmentation model[31]. The energy function is introduced with a distance regularization term to maintain the signed distance property of LSF near to the zero level-set and eliminate the need for re-initialization, as described by

$$E_{levelset}(\Phi) = \mu R_p(\Phi) + E_{ext}(\Phi) \qquad \text{Equation 3,}$$

in which μ is a positive constant. The level-set regularization term $R_p(\Phi)$ is defined with a potential function as

$$R_p(\Phi) = \int_\Omega p(\nabla\Phi)dx = \frac{1}{2}\int_\Omega (|\nabla\Phi| - 1)^2 dx \qquad \text{Equation 4.}$$

This penalty term forces the gradient magnitude of the level set function $|\nabla\Phi|$ close to 1, therefore effectively reduces the deviation of LSF from the signed distance function and ensures stable contour evolution [32].

The external constraint $E_{ext}(\Phi)$ is associated with an edge indicator function,

$$g = \frac{1}{1 + |\nabla G_\sigma * I|^2} \qquad \text{Equation 5,}$$

where G_σ is a Gaussian filter of standard deviation σ. The pre-filtering operation is necessary to smooth the image and reduce noise. Ideally, the gradient magnitude $|\nabla I|$ is maximum at the object boundaries (edge detection) and gives function g a minimum value.

$$E_{ext}(\Phi) = \lambda \int_\Omega g\delta(\Phi)|\nabla\Phi|\mathrm{dx} + \alpha \int_\Omega gH(-\Phi)\mathrm{dx} \qquad \text{Equation 6,}$$

where λ and α are both positive constants, and δ and H are the Dirac Delta and Heaviside functions respectively. The external energy term is designed to slow down the curve evolution at a location of interest, the outer lumen in our case. The first term is a line integral of function g along the zero level-set contour of Φ, which is minimized when getting closer to the object boundaries. The second term is a weighted area of Φ region that helps to speed up the motion of zero level-set contour evolution. In a

scenario that initial contour is far away from desired object location, the additional term is particularly important to shrink the contour close to the edge.

To find an optimal curve C, the objective function needs to be minimized by solving the associated Euler-Lagrange equation. According to the gradient descent concept, the zero level-set contour evolves most efficiently in the opposite direction of maximum gradient, $N = -\nabla \Phi / |\nabla \Phi|$, the steady state solution can be then solved by

$$\frac{\partial \Phi}{\partial t} = -\frac{\partial E_{levelset}}{\partial \Phi} \qquad \text{Equation 7.}$$

The Gateaux derivative of energy function gives,

$$\frac{\partial \Phi}{\partial t} = -(\mu \frac{\partial R_p}{\partial \Phi} + \frac{\partial E_{ext}}{\partial \Phi}) \qquad \text{Equation 8.}$$

$$= \mu \left[\nabla^2 \Phi - div\left(\frac{\nabla \Phi}{|\nabla \Phi|}\right)\right] + \lambda \delta(\Phi) div\left(g \frac{\nabla \Phi}{|\nabla \Phi|}\right) + \alpha g \delta(\Phi)$$

As the outer wall of a fusiform aortic aneurysm follows an elliptical contour, the GAC model was further refined with a shape constraint that preserves the circular aortic wall while isolating it from the confounding structures. The basic idea of the new energy term is to measure the area difference between an evolving shape and the desired elliptical shape. The proposed energy function is given as:

$$E_{AAA}(\Phi) = E_{levelset}(\Phi) + \beta E_{ellipse} \qquad \text{Equation 9.}$$

$$E_{ellipse} = \int_{\Omega} (\Phi - \Phi_e)^2 \delta(\Phi)\, dx \qquad \text{Equation 10.}$$

2.5 Radiomics Features Extraction

After AAA segmentation, the VOI-masked CT images were loaded into Pyradiomics for feature extraction. Four major steps were involved (i) image preprocessing; (ii)

application of enabled filters or transforms; (iii) calculation of different feature classes; iv) returning the quantitative features for analysis.

2.5.1 Preprocessing and Filtering

The image preprocessing was performed to facilitate data standardization in radiomics analysis. Normalization was firstly done to minimize the influence of contrast and brightness variations in different CT scans. The intensity of each voxel was shifted by subtracting the mean and then divided by the standard deviation. Images were then resampled into isotropic voxels with the size of $2mm^3$using a B-spline interpolator and the intensity values in the AAA volumes were discretized using a bin width of 25 Hounsfield units.

Aside from the original images, the radiomics features were also extracted from the filtered CT images. A 3D Laplacian of Gaussian (LoG) filter was used with different sigma values to detect fine to coarse textures [33]. Discrete wavelet transformations were used to focus features on different frequency ranges within the volume of interest, which yields 8 different combinations of decompositions. The image pre-processing and filtering were done by using SimpleITK [34] and PyWavelets [35].

2.5.2 Feature Classes

A total of 1130 radiomics features were extracted from each dataset, which can be divided into four categories as shown in Table 2.

Table 2. Radiomics feature class description and associated metrics [27].

Feature Class	Feature Metrics
3D Shape-Based N=14	Features calculated from VOI space using a marching cubes algorithm. No filters or transformation applied.
	Volume, surface area, sphericity, compactness, maximum 3D diameter, major axis length, etc.
1st Order Statistics N=19	Histogram-based features that describe the intensity distribution within VOI-masked images.
	Minimum, mean, maximum, skewness (asymmetry), kurtosis (flatness), entropy (randomness), 10th percentile, etc.
2nd Order Statistics N=74	Texture-based features that describe the spatial arrangement of voxel intensities within VOI-masked images.
	Gray level co-occurrence matrix (GLCM), gray level run length matrix (GLRLM), gray level size zone matrix (GLSZM), etc.
Higher Order Statistics N=1023	Features extracted from derived images using LoG with 3 sigma levels and 1 level of Wavelet decompositions.

2.6 Machine Learning Framework for AAA Classification

2.6.1 Feature Selection

A large number of radiomics features with a limited sample size could hinder the predictive power of the classification model. Feature selection is an important step to tackle the curse of dimensionality, which aims to determine a subset of the relevant features from the original ones. In the study, both non-supervised and supervised methods were evaluated.

2.6.1.1 Principal Component Analysis (PCA)

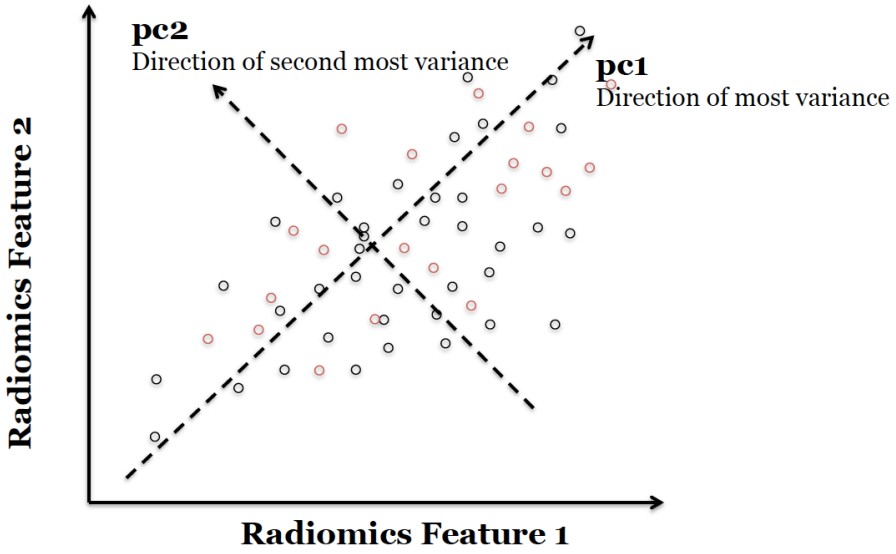

Figure 3. An illustration of principle components in a 2D radiomics features space.

PCA is an unsupervised feature selection method that uncovers low-dimensional representations of data by projecting them onto a set of independent principal components (PCs) and keeping only those that contribute most to the classification [36]. The first PC (pc1) is created by maximizing the variance of the projected points. The second PC (pc2) is selected similarly but in a direction orthogonal to pc1, so as to make

them independent to each other, as shown in Figure 3. As such, the maximum number of PCs is either the number of samples or number of features in the dataset. The PCs are considered a linear combination of the original features. A larger number of PCs means a greater variance explained by the selected features, hence the patterns of dataset. In classification problems, the optimal number of PC is determined by the prediction performance of a model.

2.6.1.2 Recursive Feature Elimination (RFE)

RFE is a supervised feature selection method that involves an iterative process to recursively remove features of low importance [37]. The algorithm fits a classification model to all features and ranks them based on their importance to the model with cross-validation (resampling). Table 3 illustrates the RFE process. Let S be a sequence of ordered numbers, which are candidate values for the number of features to retain. At each iteration in feature selection, the S_i top ranked features are retained, and the model is refit and performance is evaluated. The value of S_i with the best performance is determined and the top S_i predictors are used to fit the final model.

Table 3. RFE iterative process with subsampling and subset selections.

1	**for** *Each Resampling Iteration* **do**
2	-> Partition dataset into training and testing set via resampling.
3	-> Tune/train the estimator on training set using all features.
4	-> Predict the testing set and calculate importance rankings.
6	**for** *Each Subset Size* S_i *,* $i = 1,2,3,\ldots S$ **do**
7	-> Keep the S_i most important features.
8	-> Tune/Training the estimator on the training set using S_i features.
9	-> Predict the testing set
10	**end**
11	**end**
12	-> Calculate the performance profile over S_i using the testing set
13	-> Determine the optimal number of features
14	-> Estimate the final number of features to keep for classification model
15	-> Fit the final model based on the optimal S_i using original training set.

2.6.2 Classification Modeling

The goal of radiomics analysis is to obtain a predictive model with high accuracy and efficiency. Random forest (RF) and Logistic regression (LR) are two standard machine-learning algorithms for binary classification. RF is a non-parametric model that learns a non-linear decision boundary, whereas LR is a parametric model that learns a linear decision boundary to segment two classes, as shown in Figure 4, 5.

2.6.2.1 Random Forest (RF)

RF is an ensemble-based algorithm that operates by constructing multiple decision trees. It is built off the idea of bootstrap aggregation (random sampling with replacement) [38]. Each tree begins with a subset of randomly selected training cases. At each internal node, a randomly selected set of features is used to split the training data into child nodes that minimize class "impurity", i.e., most training cases should have the same classification within each node. The process is recursively repeated until all the samples in a node belong to the same class.

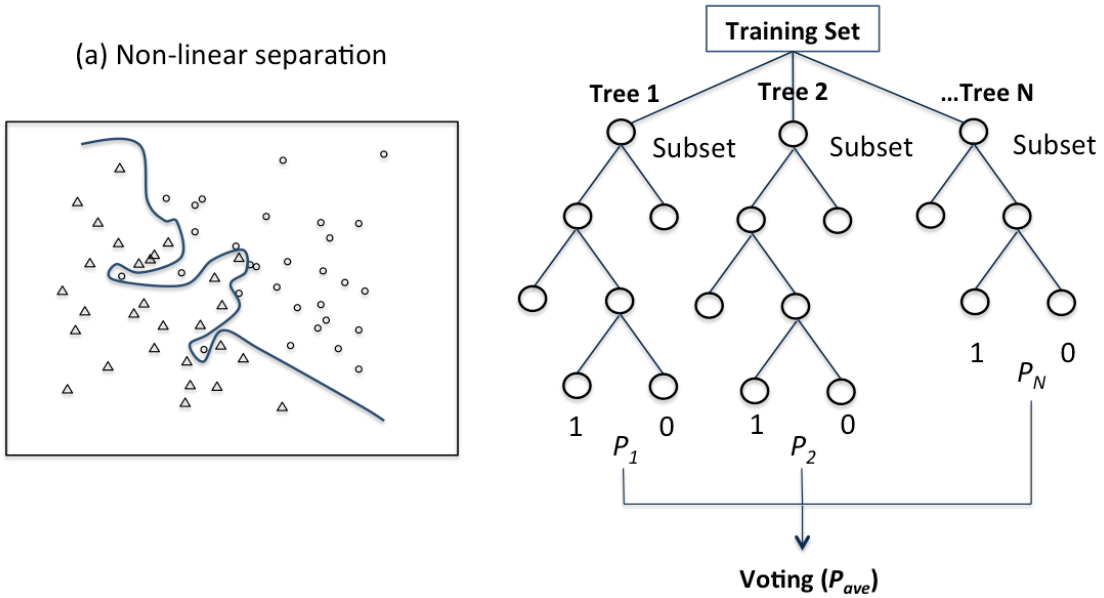

Figure 4. A non-linear separation of datasets using RF and decision tree ensembles.

The number of trees in the ensemble and the number of features selected at each node are the two main parameters of the RF algorithm. In principle, the number of trees should be sufficient to ensure that each candidate feature has the opportunity to be selected. A large number of features reduces the risk of having non-informative candidate features whereas a small number increases the chance of detecting small effects [40]. The predictive significance of each feature is determined by calculating the mean node impurity change. A larger decrease in impurity signifies feature importance. From the testing dataset, each tree gives an estimate of the probability of the class label, and the ultimate probability (*P*) for each test case is averaged over all trees.

2.6.2.2 Logistic Regression Classifier (LR)

LR is a linear model that describes the relationship between multiple independent variables (features) and a dependent response variable (class) [39]. It transforms the output using a logistic sigmoid function, as shown in Figure 4, to return the probabilities of the label class. The sigmoid function has the advantage to take a real-value number and map it into the range [0, 1] for continuous probabilities.

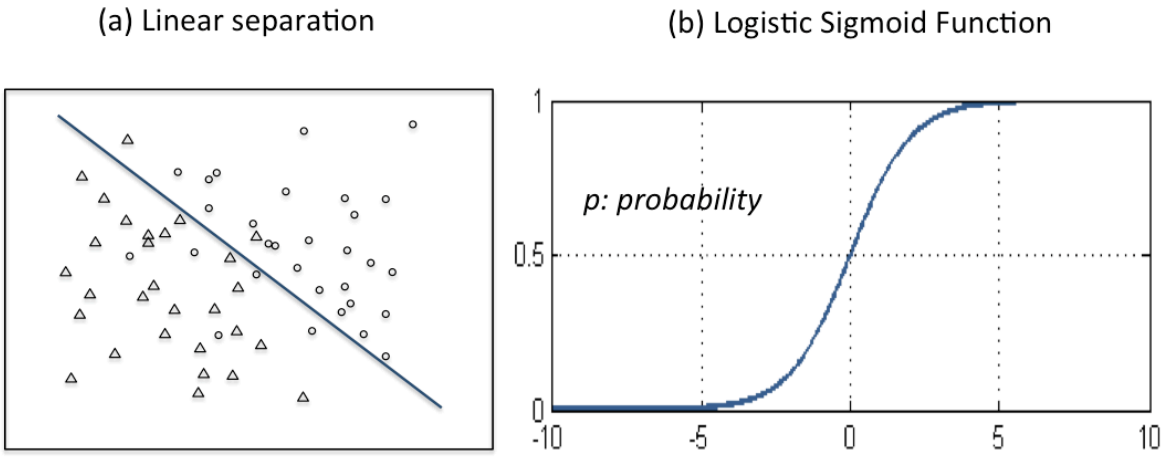

Figure 5. A linear separation of datasets using LR and the logistic sigmoid function.

Let Y denote the class variable and $x_1, \ldots, x_n$ the feature variables. The LR model links the estimated probability $P\ (Y = 1 \mid x_1, x_2, \ldots, x_n)$ to $x_1, x_2, \ldots, x_n$ through

$$P(Y = 1|x_1, x_2, \ldots, x_n) = \frac{e^{\beta_0+\beta_1(x_1)+\beta_2(x_2)+\cdots\beta_n(x_n)}}{1 + e^{\beta_0+\beta_1(x_1)+\beta_2(x_2)+\cdots\beta_n(x_n)}} \qquad \text{Equation 11,}$$

where $\beta_0, \beta_1, \ldots, \beta_n$ are the regression coefficients determined by the maximum-likelihood estimation (MLE)[39].

2.6.3 Implementation of Radiomics Classification Model

In our study, the feature selection and classification algorithms described above were provided by the scikit-learn Python package (v0.21.3). The hyper-parameters were fine-tuned using a grid search algorithm that tests all possible combinations of parameter values for feature selection and classifiers. The best combination was retained for the prediction evaluation. To avoid overfitting, we used k-fold cross-validation with k=5. The dataset is partitioned into 5 subsets of equal size (fold). For 5 iterations, one fold was chosen as the testing set, while the remaining folds were used for training, and the overall performance of the classifier was averaged over all iterations.

2.7 Classification Performance Evaluation

Classifier performance was first evaluated by the best prediction accuracy, defined as the average percentage of correct classification after 5-fold cross-validation. Subsequently, the receiver-operating characteristic (ROC) was derived for each testing fold based on the output probabilities from the classifier. The mean area under curve (AUC) value was used for the comparisons between different feature selection and classification combinations.

To further investigate the predictive performance of radiomics features as compared to the clinical standard, the conventional risk factors (clinical feature), including D_{max},

age, BMI, smoking (1/0), HT (1/0) and DM (1/0), were also fed into LR and RF classifiers to identify the fast-growing AAA. The class probabilities derived from the classifiers were subsequently used for ROC analysis in STATA (release 16) without cross-validation. A DeLong test [40] was performed to evaluate the statistical significance of the difference between the areas under 2 dependent ROC curves. Using the univariate and multivariate logistic regression analyses, the association of radiomics features with the prediction outcome was summarized. With 95% confidence interval, the predictive values were assessed using odds ratios (ORs).

3. RESULTS

3.1 Quantitative Evaluation of Segmentation Results

From 22 randomly selected datasets, the segmentation results were evaluated quantitatively using the dice similarity coefficient (DSC) and volume difference (VD). Comparisons were made between the semi-automatic segmentation (V_1) and the ground-truth standard (V_2) obtained by manual segmentation from a trained radiologist.

The DSC is defined as the volumetric overlap index that measures the similarity between two sets of segmentations. Mathematically, it is represented as

$$DSC = \frac{2|V_1 \cap V_2|}{|V_1| + |V_2|} \qquad \text{Equation 12,}$$

in which $\cap$ is an intersection operator and $|\ |$ denotes the total number of overlapping pixels. Given a range from 0 to 1, the higher the fraction index is, the greater the similarity between semi-automatic and manual segmentations.

The VD is given by

$$VD = \frac{|V_1 - V_2|}{\frac{1}{2}(|V_1| + |V_2|)} \qquad \text{Equation 13.}$$

Table 4. Quantitative evaluation of the semi-automatic segmentation results.

	%DSC ± SD	%VD ± SD
Outer Wall	92.9 ± 2.9	7.0 ± 6.9

Using linear regression, a trend-line was fit to investigate the correlations between the volumes calculated from V_1 and V_2. A correlation coefficient of 0.995 was achieved. A Bland-Altman analysis was also conducted to assess the agreement. No bias was observed as shown in Figure 6.

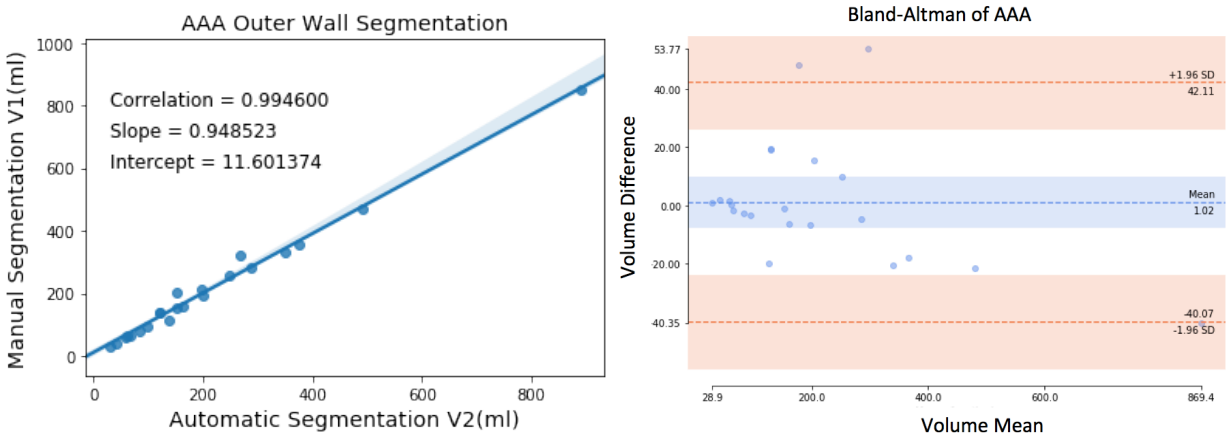

Figure 6. A linear trend-line that fits the pixel volume from semi-automatic and manual segmentation results, and a Bland-Altman plot that compares the volume difference and mean between them.

3.2 Prediction Performance

Table 5 shows the prediction accuracies of two classifiers combined with two feature selection algorithms under 5-fold cross-validation. With PCA, the first 25 and 15 PCs were selected for RF and LR classifiers respectively, based on grid searching. By preserving a large number of PCs, RF yielded better classification accuracy than LR. With RFE, 10 features were selected and the same prediction accuracies were achieved for RF and LR classifiers.

Table 5. The prediction accuracies of classifiers combined with feature selection algorithms.

	RF		**LR**	
	Accuracy (%)	**No. of Features**	**Accuracy (%)**	**No.of Features**
PCA	73.8%	25	69.0%	15
RFE	75.0%	10	75.0%	10

The ROC results shown in Figure 7 indicate that RFE has significantly higher mean AUC values (0.8 and 0.8) than that of PCA (0.62 and 0.58) for both classifiers. In the 5 iterations, the RFE yielded a relatively smaller variance indicating its potential to select more stable features for the classification models.

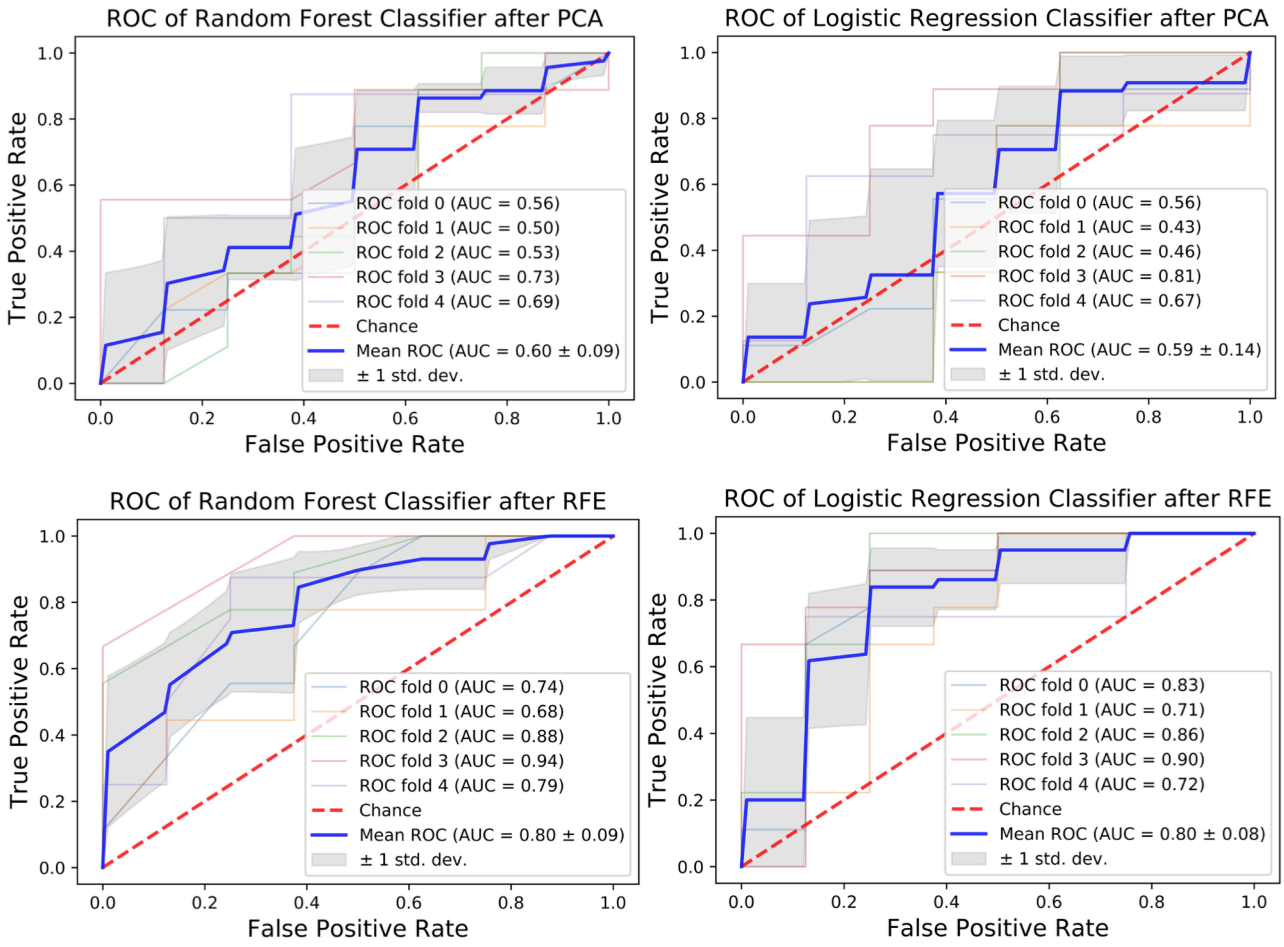

Figure 7. The ROC responses for four feature selection and classifier combinations. Each graph presents 5-fold cross-validation, with the mean ROC shown in blue and variance shaded in grey.

3.3 Selected Feature Importance

From RFE, 10 features were shortlisted and ranked according to their feature importance in RF, as shown in Figure 7 (see Table 2 for brief feature descriptions). The most important feature was a first-order statistics (wavelet-LLL_firstorder_10Percentile) extracted from the wavelet-filtered image. The axial maximum diameter (original_shape_Maximum2DDiameterRow) was ranked the 3rd most important feature. Majority of the other features were second-order statistics (e.g.,

glcm, ngtdm) that characterize the textures (gray-scale patterns) of AAA in the CT images.

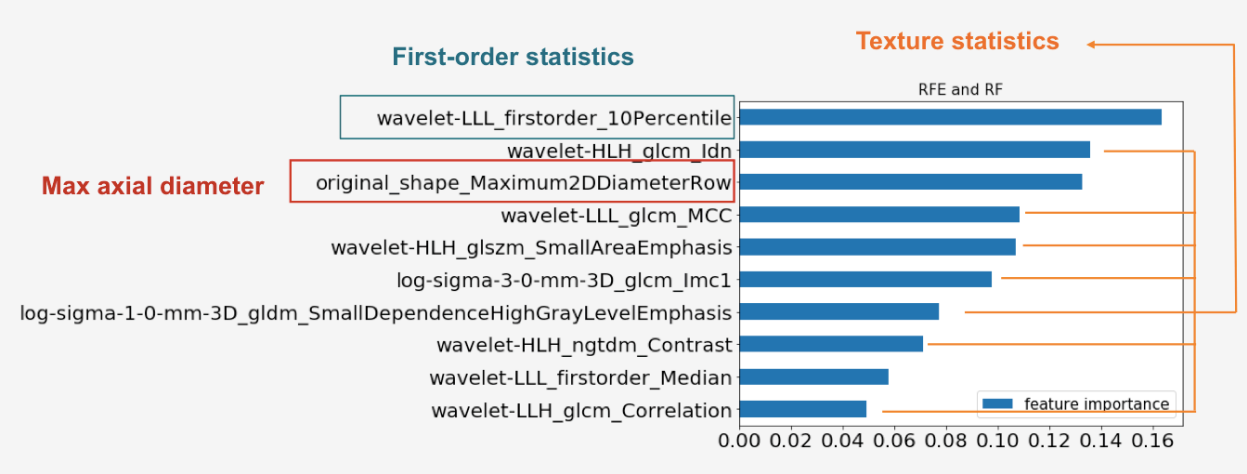

Figure 8. The importance ranking of selected radiomics features from RFE in RF.

3.4 Comparison of Radiomics Features and Clinical Features

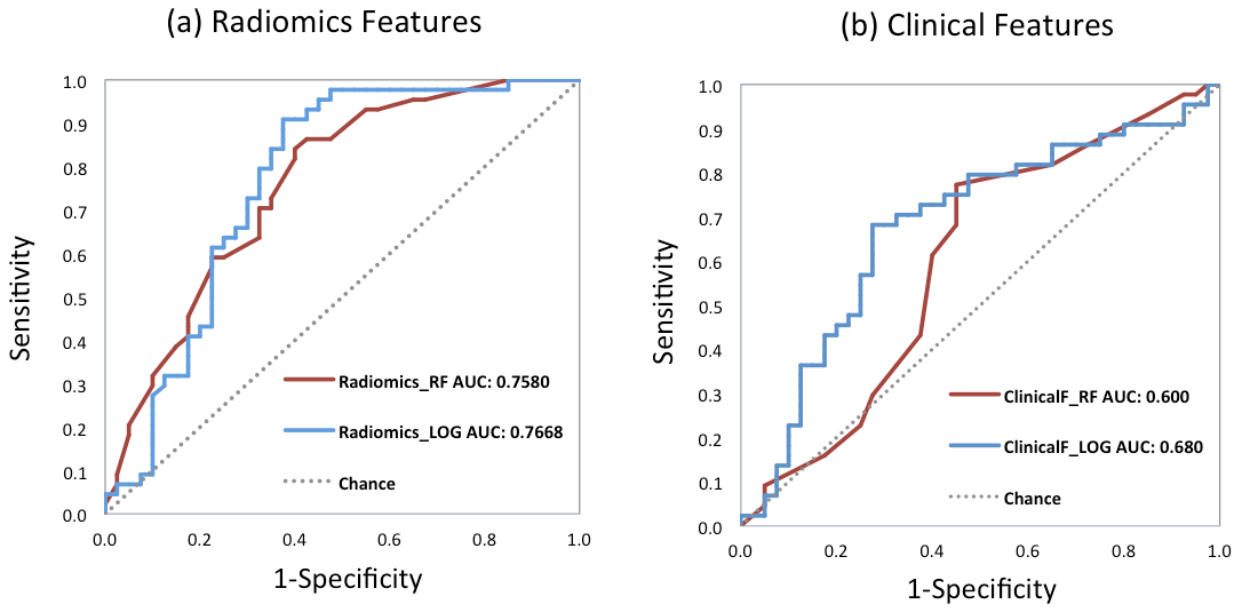

Figure 9. The ROC responses for two classifiers with (a) radiomics features and (b) clinical features, plotted in STATA without cross-validation.

As seen from Figure 9, the radiomics features gave relatively higher AUC values as compared to the clinical features in both classifiers. This is more evident in the high false positive rate range, where the steepness of ROC is smaller for radiomics features.

Table 6. The DeLong test results showing the statistical significance of difference between AUCs of two classifiers and features types.

	Random Forest AUC [95% CI]	**Logistic Regression** AUC [95% CI]	***p-value*** (Classifiers)
Radiomics Features	0.758 [0.654, 0.862]	0.767 [0.658, 0.876]	0.796
Clinical Features	0.600 [0.474, 0.725]	0.680 [0.561, 0.798]	0.0504
p-value (Features)	0.017	0.0611	

No significant difference was observed in the ROC from the two classifiers. This is shown in the DeLong test results in Table 6, in which the AUC performance between two classifiers yielded p-values of 0.796 and 0.0504 for radiomics features and clinical features respectively. However, by looking at the last row that compares the AUC performance between the radiomics features and clinical features, it yielded a p-value of 0.017 and 0.0611 for RF and LR classifiers, respectively. This indicated that RF classifier implemented with radiomics features outperforms the clinical feature in identifying AAA with rapid expansion.

In univariate analysis, the D_{max} (OR:2.33, p=0.003, 95% CI:1.336 – 4.067), clinical features (OR:4.72, p=0.1, 95% CI:0.714 – 31.237), and radiomics features (OR:63.20, p<0.001, 95% CI:8.151 – 490.055) were all found to be positively associated with fast growth rate. Nonetheless, the results from the multivariate analysis showed that only RF with radiomics features (OR: 38.105; 95% CI:3.489 - 416.200) were independently

associated with rapid expansion. The goodness of fit for the model was acceptable (N= 84, $\chi 2 = 21.3$, $p < 0.001$).

4. DISCUSSION

This study set out with an aim to determine the feasibility of radiomics in identifying AAA with a fast growth rate (≥0.3cm/year) using CT images. Accurate and fast segmentation is critical for radiomics analysis. While manual segmentation is time-consuming, a semi-automatic AAA segmentation algorithm was developed to process a large-scale CT dataset. A total of 1130 radiomics features were extracted from the segmented VOI. In an ML framework built with two feature selection algorithms and two classifiers, 10 radiomics features were selected for the AAA classification. The ROC evaluation has shown that the radiomics classification model outperformed the conventional risk factors (clinical features) in distinguishing the fast-growing and stable AAA. Further multivariate analysis suggested that this model can be used as an independent, yet strong predictor for fast AAA growth rate.

4.1 Segmentations

The semi-automatic segmentation algorithm was developed with a GAC model, in which the contour was formulated as the evolution of a level-set function associated with energy constraints. Compared to previous studies [41-43], in which the reported accuracies of segmentation for DCS and VD were in the range of 90.1%-95.3% and 6.2%-16%, our method produced comparable results (DSC = 92.9 ± 2.9 %, VD = 7.0 ± 6.9% (Table 4)).

Despite the accurate segmentation outcome, the method has some limitations. One limitation is that it requires a manual reference contour to initiate the automatic

process. This inevitably makes the first step user-dependent. Although the manual contouring of the initial slice has been demonstrated to be reproducible in a recent study[32], an automatic detection mechanism is desirable to increase the efficiency of segmentation. Another limitation lies in the mechanism of automatic detection, in which the contour of a new slice relies on the segmentation result from the previous slice. An accumulative error could occur and lead to inaccurate results. Since our method allows for a user-interactive environment, the error can be mitigated by necessary manual intervention.

4.2 The ML Framework

Two feature selection algorithms, PCA and RFE, were implemented with two classifiers to investigate the predictive power of radiomics features. Overall, RFE yielded higher prediction accuracies and better ROC performance as compared to PCA (Table 6). One reason is that RFE specifically removes irrelevant and unstable features based on the supervision, while PCA is an unsupervised feature transformation not tailored to any classifier. Moreover, the PCs were selected so that they can describe the majority of the variance in the dataset and use these as inputs to the ML classifiers. However, this is based on the assumption that the PCs with the highest variance also contain the most information to separate a dataset by its labels, meaning PCA-based classification may provide subpar results if this assumption is incorrect.

RFE should be applied with caution since it is prone to over-fitting. Unlike PCA, it is highly selective and ignores the effect of interaction among features [44]. Hence, it might not work well for new samples and classifiers.

With a greater number of PCs, RF gave higher prediction accuracy over LR, indicating that RF copes better with large feature dimensions [21]. One explanation is

that RF is immune to redundant features or features with high correlation, and performs well with outliers [33, 45]. On the contrary, LR requires each feature to be independent. Highly-correlated (dependent) features will dilute the individual effect of the feature in the regression model [46].

With the same selected features from RFE, RF and LR classifiers performed similarly to each other in ROC ($p>0.05$), though the current sample size is limited. Given its simplicity and interpretability, LR is a good classification model for linear separations that can be implemented with less effort. However, the feature selection method should be incorporated to remove collinearity in the dataset beforehand. For datasets with more complicated and non-linearly distributed radiomics features, RF will probably perform better.

4.3 Radiomics Features versus Clinical Features

Among the 10 selected features, the axial maximum diameter (original_shape_Maximum2DDiameterRow) of AAA is ranked as the 3rd most important feature that contributes to the classification model. This is consistent with the fact the D_{max} is the primary and most significant clinical predictor for AAA fast growth rate. However, the other geometric metrics like volume or surface area do not stand out as significant features for fast growth rate, which contradicts the results from previous literature [20]. It was suspected that different segmentation strategies, particularly, in defining the aneurysmal region, could lead to significant variability in the quantification of geometric metrics. In this regard, a more consistent or precise definition of the AAA region is necessary to further explore the use of geometric indices. To our surprise, the top-ranked feature (wavelet-LLL_firstorder_10Percentile) was a simple first-order statistic that describes the 10th percentile of the intensity distribution in AAA after the

wavelet transformation. This feature is speculated to be associated with a higher amount of intraluminal thrombus, which leads to a lower intensity in the aneurysm volume. This is corroborated further by the previous studies, which suggests that ILT volume is an independent predictor for rapid AAA expansion [15, 16]. The majority of other features were texture metrics that describe the heterogeneity or gray-scale patterns in the AAA volumes. The lumen contrast heterogeneity was quantified and determined to be independently associated with rapid growth rate [47]. However, currently, the association between texture statistics and physical AAA characteristics is nebulous.

Based on the AUC values, the classification model built with radiomics features outperformed the model built with clinical features. The multivariate logistic regression analysis shows that the radiomics classification model was independently associated with the prediction outcome (OR: 63.20, $p<0.001$, 95% CI: 8.151 – 490.055). Since D_{max} is included in radiomics features, it makes sense that radiomics-based ML classification is superior to conventional classification based on D_{max} alone. Hence, we conclude that radiomics combined with the proper ML framework do have predictive value for AAA progression.

4.4 Limitations

In radiomics, the power of classification model is dependent on the sample size, quality of the database, segmentation algorithm, and feature extraction settings. Some special considerations deserve further attention. This study was limited by the small cohort size of 84 patients. Furthermore, it also lacked an independent validation of the stability and robustness of the selected radiomics features. When medical images are analyzed numerically to extract meaningful data, variations in image processing and segmentation can introduce changes that are not due to underlying biologic effects. Bias

and variance may come from both manual and semi-automatic segmentation, resulting in different radiomics features extracted. Using multiple methods for segmentation can help to understand the extent of this bias. Several studies have investigated feature stability through a test-retest process on medical images [48], or by comparing the results obtained with different acquisition settings and processing algorithms [49]. Other than RF classification algorithms, methods like naive Bayesian and K-nearest neighbors have also been incorporated in radiomics feature analysis. The performance and robustness await further evaluations in our application. Moreover, a more stringent feature selection scheme can be considered. A repeated cross-validation is necessary to examine the efficiency of classifier and its stability in selecting important features. For this longitudinal study, other confounding factors like medications, other disease conditions (high cholesterol, inflammatory state), and even sugar intake and lifestyle may have causal effect on the growth of AAA within the 6-month timeline. Hence, further considerations should be given to the patient-specific characteristics.

5. CONCLUSION

As a clinical diagnosis made with visual assessment can be affected by inter-reader variations and clinical experience, a reproducible and accurate method for monitoring AAA progression is unmet. Given wide applications in oncology, radiomics is believed to be a potential tool to assist diagnosis in AAA. This is supported by our pioneering results. With an expanded cohort study and improved analysis pipeline, radiomics could offer a more comprehensive characterization of AAA growth.

6. REFERENCES

[1] Y. Kumar, K. Hooda, S. Li, P. Goyal, N. Gupta, and M. Adeb, "Abdominal aortic aneurysm: pictorial review of common appearances and complications," *Ann Transl Med,* vol. 5, p. 256, Jun 2017.

[2] V. A. S. S. Chichester Aneurysm Screening Group, "A comparative study of the prevalence of abdominal aortic aneurysms in the United Kingdom, Denmark, and Australia," *J Med Screen,* 2001.

[3] P. E. Norman and J. T. Powell, "Abdominal aortic aneurysm: the prognosis in women is worse than in men," *Circulation,* vol. 115, pp. 2865-9, Jun 5 2007.

[4] E. L. Chaikof, R. L. Dalman, M. K. Eskandari, B. M. Jackson, W. A. Lee, M. A. Mansour*, et al.*, "The Society for Vascular Surgery practice guidelines on the care of patients with an abdominal aortic aneurysm," *J Vasc Surg,* vol. 67, pp. 2-77 e2, Jan 2018.

[5] A. Mathur, V. Mohan, D. Ameta, B. Gaurav, and P. Haranahalli, "Aortic aneurysm," *J Transl Int Med,* vol. 4, pp. 35-41, Apr 1 2016.

[6] A. T. Hirsch, Z. J. Haskal, N. R. Hertzer, C. W. Bakal, M. A. Creager, J. L. Halperin*, et al.*, "ACC/AHA 2005 Practice Guidelines for the management of patients with peripheral arterial disease (lower extremity, renal, mesenteric, and abdominal aortic): a collaborative report from the American Association for Vascular Surgery/Society for Vascular Surgery, Society for Cardiovascular Angiography and Interventions, Society for Vascular Medicine and Biology, Society of Interventional Radiology, and the ACC/AHA Task Force on Practice Guidelines (Writing Committee to Develop Guidelines for the Management of Patients With Peripheral Arterial Disease): endorsed by the American Association

of Cardiovascular and Pulmonary Rehabilitation; National Heart, Lung, and Blood Institute; Society for Vascular Nursing; TransAtlantic Inter-Society Consensus; and Vascular Disease Foundation," *Circulation,* vol. 113, pp. e463-654, Mar 21 2006.

[7] M. J. T. Sweeting, S.G. Brown, L.C. Powell, J.T., "Meta-analysis of individual patient data to examine factors affecting growth and rupture of small abdominal aortic aneurysm," *Br J Surg,* vol. 99, pp. 655-665, 2012.

[8] S. G. Thompson, L. C. Brown, M. J. Sweeting, M. J. Bown, L. G. Kim, M. J. Glover*, et al.*, "Systematic review and meta-analysis of the growth and rupture rates of small abdominal aortic aneurysms: implications for surveillance intervals and their cost-effectiveness," *Health Technol Assess,* vol. 17, pp. 1-118, Sep 2013.

[9] A. N. Dmitry Rakita, John J.Hines, David N Siegel, Barak Friedman, "Spectrum of CT Findings in Rupture and Impending Rupture of Abdominal Aortic Aneurysms," *RadioGraphics,* vol. , 2007.

[10] R. H. C. Cary L. Siegel, Melvin Korobkin, Michael B. Alpern, Daniel L. Courneya, Richard A. Leder, "Abdominal Aortic Aneurysm Morphology CT Features in Patients with Rupture and Nonrupture," *AJR* vol. 163, pp. 1123-1129, 1994.

[11] P. J. C. Thorsten A. Bley, Scott B. Reeder, Christopher J.Francois, Kazuhiko Shinki, Girma Tefera, Frank N. Ranallo, Thomas M. Grist, Myron Pozniak, "Endovascular Abdominal Aortic Aneurysm Repair- Nonenhanced Volumetric CT for Follow-up," *Radiology* vol. 253, 2009.

[12] A. Parr, M. McCann, B. Bradshaw, A. Shahzad, P. Buttner, and J. Golledge, "Thrombus volume is associated with cardiovascular events and aneurysm

growth in patients who have abdominal aortic aneurysms," *J Vasc Surg,* vol. 53, pp. 28-35, Jan 2011.

[13] W. J. Q. William R.Mower, Sanjiv S. Gambhir, , "Effect of intraluminal thrombus on abdominal aortic aneurysm wall stress," *The Society for Vascular Surgery and International Society for Cardiovascular Surgery,* vol. 26, 1997.

[14] D. H. Wang, M. S. Makaroun, M. W. Webster, and D. A. Vorp, "Effect of intraluminal thrombus on wall stress in patient-specific models of abdominal aortic aneurysm," *J Vasc Surg,* vol. 36, pp. 598-604, Sep 2002.

[15] B. K. Carrell TW, Booth NA, Humphries J, Smith A., "Intraluminal thrombus enhances proteolysis in abdominal aortic aneurysms.," *Vascular,* vol. 14, 2006.

[16] M. N. A. A. R. Junaid A.Khan, F.A.K Mazari, Y.Shahin, G.Smith, L.Madden, M.J. Fagan, J. Greenman, P.T. McCollum, I.C. Chetter, "Intraluminal Thrombus has a Selective Influence on Matrix Metalloproteinases and Their Inhibitors (Tissue Inhibitors of Matrix Metalloproteinases) in the Wall of Abdominal Aortic Aneurysms," *Annals of Vascular Surgery,* vol. 26, pp. 322-329, 2012.

[17] R. V. Buijs, T. P. Willems, R. A. Tio, H. H. Boersma, I. F. Tielliu, R. H. Slart*, et al.*, "Calcification as a risk factor for rupture of abdominal aortic aneurysm," *Eur J Vasc Endovasc Surg,* vol. 46, pp. 542-8, Nov 2013.

[18] Z. Y. Li, U. K.-I. J, T. Y. Tang, E. Soh, T. C. See, and J. H. Gillard, "Impact of calcification and intraluminal thrombus on the computed wall stresses of abdominal aortic aneurysm," *J Vasc Surg,* vol. 47, pp. 928-35, May 2008.

[19] J. Shum, G. Martufi, E. Di Martino, C. B. Washington, J. Grisafi, S. C. Muluk*, et al.*, "Quantitative assessment of abdominal aortic aneurysm geometry," *Ann Biomed Eng,* vol. 39, pp. 277-86, Jan 2011.

[20] S. A. Parikh, R. Gomez, M. Thirugnanasambandam, S. S. Chauhan, V. De Oliveira, S. C. Muluk*, et al.*, "Decision Tree Based Classification of Abdominal Aortic Aneurysms Using Geometry Quantification Measures," *Ann Biomed Eng,* vol. 46, pp. 2135-2147, Dec 2018.

[21] Y. Zhang, A. Oikonomou, A. Wong, M. A. Haider, and F. Khalvati, "Radiomics-based Prognosis Analysis for Non-Small Cell Lung Cancer," *Sci Rep,* vol. 7, p. 46349, Apr 18 2017.

[22] P. E. K. Robert J.Gillies, Hedvlg Hricak, "Radiomics: Images Are More than Pictures, They Are Data," *Radiology,* vol. vol. 278, pp. 563-77, 2016.

[23] H. Cao, S. Bernard, R. Sabourin, and L. Heutte, "Random forest dissimilarity based multi-view learning for Radiomics application," *Pattern Recognition,* vol. 88, pp. 185-197, 2019.

[24] G. Garcia, J. Maiora, A. Tapia, and M. De Blas, "Evaluation of texture for classification of abdominal aortic aneurysm after endovascular repair," *J Digit Imaging,* vol. 25, pp. 369-76, Jun 2012.

[25] J. H. F. R. Carl W. Kotzea, Balaji Ganeshana, Leon J. Menezesa, Jocelyn Brookesa, Obiekezie Agua, Syed W. Yusufc, Ashley M. Grovesa,∗,'Correspondence information about the author Ashley M. GrovesEmail the author Ashley M. Groves, "CT signal heterogeneity of abdominal aortic aneurysm as a possible predictive biomarker for expansion," *Atherosclerosis,* vol. 233, pp. 510-517, 2014.

[26] M. Kolossvary, J. Karady, B. Szilveszter, P. Kitslaar, U. Hoffmann, B. Merkely*, et al.*, "Radiomic Features Are Superior to Conventional Quantitative Computed

Tomographic Metrics to Identify Coronary Plaques With Napkin-Ring Sign," *Circ Cardiovasc Imaging,* vol. 10, Dec 2017.

[27] J. J. M. van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan*, et al.*, "Computational Radiomics System to Decode the Radiographic Phenotype," *Cancer Res,* vol. 77, pp. e104-e107, Nov 1 2017.

[28] P. a. al, "Scikit-learn: Machine Learning in Python," *JMLR* vol. 12, 2011.

[29] R. D. Nikos Paragios, "Geodesic Active Contours and Level Sets for the Detection and Tracking of Moving Objects," *IEEE,* vol. 22, 2000.

[30] R. F. Stanley Osher, "LevelSetMethods," *Applied Mathematical Sciences,* vol. 153, 2000.

[31] J. A. S. Ravikanth Malladi, Baba C. Vemuri, "Shape Modeling with Front Propagation- A Level Set Approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 17, pp. 158-175, 1995.

[32] Y. Wang, Y. Zhang, W. Xuan, E. Kao, P. Cao, B. Tian*, et al.*, "Fully automatic segmentation of 4D MRI for cardiac functional measurements," *Med Phys,* vol. 46, pp. 180-189, Jan 2019.

[33] Z. Hou, Y. Yang, S. Li, J. Yan, W. Ren, J. Liu*, et al.*, "Radiomic analysis using contrast-enhanced CT: predict treatment response to pulsed low dose rate radiotherapy in gastric carcinoma with abdominal cavity metastasis," *Quant Imaging Med Surg,* vol. 8, pp. 410-420, May 2018.

[34] B. C. Lowekamp, D. T. Chen, L. Ibanez, and D. Blezek, "The Design of SimpleITK," *Front Neuroinform,* vol. 7, p. 45, 2013.

[35] R. G. Gregory R. Lee, Filip Waselewski, Kai Wohlfahrt, and Aaron O'Leary, "PyWavelets- A Python package for wavelet analysis," *JOSS,* 2019.

[36] J. Lever, M. Krzywinski, and N. Altman, "Principal component analysis," *Nature Methods,* vol. 14, pp. 641-642, 2017.

[37] J. W. SABELLE GUYON, STEPHEN BARNHILL, "Gene Selection for Cancer Classification using Support Vector Machines," *SpringerLink, Kluwer Academic Publishers,* 2002.

[38] X. Zhu, X. Du, M. Kerich, F. W. Lohoff, and R. Momenan, "Random forest based classification of alcohol dependence patients and healthy controls using resting state MRI," *Neurosci Lett,* vol. 676, pp. 27-33, May 29 2018.

[39] R. Couronne, P. Probst, and A. L. Boulesteix, "Random forest versus logistic regression: a large-scale benchmark experiment," *BMC Bioinformatics,* vol. 19, p. 270, Jul 17 2018.

[40] E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach," *Biometrics* vol. 44, 1988.

[41] T. Siriapisith, Worapan Kusakunniran, Peter Haddawy, "3D Segmentation of Exterior Wall Surface of Abdominal Aortic Aneurysm from CT Images Using Variable Neighborhood Search," *Computers in Biology and Medicine,* vol. 107.

[42] F. Zhuge, et al, "An Abdominal Aortic Aneurysm Segmentation Method: Level Set with Region and Statistical Information.," *Medical Physics,* vol. 33, pp. 1440-1453, 2006.

[43] K. López-Linares, et al, "Fully Automatic Detection and Segmentation of Abdominal Aortic Thrombus in Post-Operative CTA Images Using Deep Convolutional Neural Networks.," *Medical Image Analysis,* vol. 46, pp. 202-214, 2018.

[44] J. G. B. Dy, C. E. , "Feature Selection for Unsupervised Learning," *J. Mach. Learn.,* vol. 5, pp. 845-889.

[45] H. H. Cho, S. H. Lee, J. Kim, and H. Park, "Classification of the glioma grading using radiomics analysis," *PeerJ,* vol. 6, p. e5982, 2018.

[46] P. e. a. Ranganathan, "Common pitfalls in statistical analysis: Logistic regression," *Perspectives in clinical research* vol. 8.3, pp. 148-151, 2017.

[47] A. Aghayev, et a, "Common First-Pass CT Angiography Findings Associated With Rapid Growth Rate in Abdominal Aorta Aneurysms Between 3 and 5 Cm in Largest Diameter," *American Journal of Roentgenology,* vol. 210, pp. 431-437, 2018.

[48] L. R. Van Timmeren JE, "Test-retest data for radiomic feature stability analysis: generalizable or study-specific?," *Tomography,* vol. 2, pp. 361-365, 2016.

[49] M. A. Solomon Justin, Nelson Rendon C., Roy Choudhury Kingshuk, Samei Ehsan, "Quantitative Features of Liver Lesions, Lung Nodules, and Renal Stones at Multi–Detector Row CT Examinations: Dependency on Radiation Dose and Reconstruction Algorithm," *Radiology,* vol. 279, pp. 185-194, 2016.

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature Tei Liong Date Sep 08 2019